

ML2020SPRING HW11 Report

學號：R08946015 系級：資料科學碩一 姓名：陳鈞廷

1. 訓練一個 model

- 請描述你使用的 model(可以是 baseline model)。包含 generator 和 discriminator 的 model architecture、loss function、optimizer 參數、以及訓練 step 數(或是 epoch 數)。
- 請畫出至少 16 張 model 生成的圖片。

我使用助教的 sample code 中的 baseline model 作為這題的 model，generator 的 input 是 dimension 為 100 的 vector，接著 1 層 fully connected layer，然後接續著 3 層 deconvolution layer 來做 upsampling 的動作，最後輸出生成的圖片；discriminator 主要包含了 5 層 convolution layer，最後輸出圖片的真假二元類別。Loss function 主要是計算圖片的 discriminator 預測結果與是否真假的 label 做 binary cross entropy，discriminator 的 loss function 是將生成圖片以及真實圖片得到的 binary cross entropy 相加起來除以 2；而 generator 的目的是希望生成圖片誤導 discriminator，換句話說就是讓 discriminator 的生成圖片的 loss 變大，因此 generator 的 loss 是生成圖片以及標籤 1 (real label) 的 binary cross entropy。

在模型的訓練上使用 Adam 這個 optimizer，learning rate 設定成 10^{-4} (設定成 10^{-3} 則無法收斂)，batch size 設定成 64，因此每個 epoch 會有 1115 個 step，總共訓練 10 個 epoch。下圖是第 10 個 epoch 的 model 的生成圖片。



2. 訓練一個 model

- 同 1.a，請描述你選擇的 model，包含 generator 和 discriminator 的 model architecture、loss function、optimizer 參數、及訓練 step 數(或是 epoch 數)。
- 和 1.b 一樣，就你選擇的 model，畫出至少 16 張 model 生成的圖片
- 請簡單探討你在 1. 使用的 model 和 2. 使用的 model，他們分別有何性質，描述你觀察到的異同。

在這題我使用了 WGAN 的架構，WGAN 的模型架構與 DCGAN 相似，主要有 4 點差異

1. 移除掉了 discriminator 輸出的 sigmoid，讓 discriminator 負責 regression task。
2. Loss function 移除掉了 binary cross entropy 的 log 運算元

3. 每次 training 的時候做 clamp 的動作限制住 parameter 的 L_∞ norm (絕對值中的最大值)
4. 不使用 momentum-based 的 optimizer，改用 SGD 或是 RMSProp。

因此在 model 的訓練上，我使用 RMSProp，其他 learning rate、batch size 與 epoch 數的設定與上題 DCGAN 相同。下圖是第 10 個 epoch 的 model 的生成圖片。



在實驗的時候，我觀察到 DCGAN 和 WGAN 最大的差異是在 learning 的設定，如果 learning rate 設定成 10^{-3} ，DCGAN 則無法收斂，會生成無意義的空白圖片；但 WGAN 則可以成功訓練起來，推測是因為 parameter 的 clamp 以及應用了 RMSprop 的 optimizer。

3. 請訓練一個會導致 mode collapse 的 model。

- 同 1.a，請描述你選擇的 model，包含 generator 和 discriminator 的 model architecture、loss function、optimizer 參數、及訓練 step 數(或是 epoch 數)。
- 請畫出至少 16 張 model 生成且具有 mode collapse 現象的圖片。
- 在不改變 optimizer 和訓練 step 數的情況下，請嘗試使用一些方法來減緩 mode collapse。說明你嘗試了哪些方法，請至少舉出一種成功改善的方法，若有其它失敗的方法也可以記錄下來。

這題我使用了跟第一題一樣的 baseline model，model architecture、loss function、optimizer 和 batch size 都跟第一題相同，唯一更動的是將 epoch 增加到 40，當 model 訓練到第 30 個 epoch 的時候，generator 的 loss 開始爆增，生成出來的圖變成完全不像人臉圖，過了 3 個 epoch 後，當 loss 下降，就出現了 mode collapse 的現象，生成出來的圖有髮型相似的特性，如下面的圖片所示：



而我改善 mode collapse 的方法是將模型換成 WGAN，就算 train 到第 40 個 epoch 仍然沒有出現 mode collapse 的現象，因此 WGAN 是一個有效避免 mode collapse 的方法。